Supplementary Figure Legend

Supplementary Fig. 1. Confidence judgment and reaction time (RT) results. **(A-C)** Confidence judgment results. **(A)** Behavioral subjects. **(B)** EEG subjects. **(C)** fMRI subjects. **(A1)** Explicit confidence ratings showed highest confidence for unambiguous faces and lowest for the most ambiguous (50% fear/50% happy) faces. **(A3, B1, C1)** The reaction time (RT; relative to stimulus onset) for the fear/happy decision can be considered as an implicit measure of confidence because it showed a similar pattern as the explicit ratings. When we grouped the 7 morph levels into 3 ambiguity levels (unambiguous, intermediate, and high), both explicit **(A2)** and implicit **(A4, B2, C2)** confidence measures varied systematically as a function of ambiguity. In addition, subjects judged facial emotions faster when they subsequently indicated higher confidence **(A5)**, and they tended to report confidence faster for higher confidence **(A6)**. Their RT to report confidence (relative to the onset of the confidence rating message) also varied as a function of morph levels **(A7)** and ambiguity levels **(A8)**. Note that the behavioral patterns of all three subject groups were comparable. Error bars denote one SEM across subjects.

Supplementary Fig. 2. Control LPP analyses. **(A, B)** The LPP still differentiated levels of ambiguity when using the average of 6 electrodes surrounding Pz (CP3, CPz, CP4, P3, Pz and P4) for both mean ($F(2,44)=12.7$, $P=4.31\times10^{-5}$, $\eta_p^2=0.37$) and peak amplitude ($F(2,44)=10.7$, $P=1.59\times10^{-4}$, $\eta_p^2=0.33$). **(A)** ERP. **(B)** Mean LPP amplitude. **(C-E)** A speeded version of the main task replicated the results (N=18). **(C)** RT. Error bars denote one SEM across subjects. **: $P<0.01$. **(D)** ERP. **(E)** Mean LPP amplitude. Conventions as **Fig. 2D, E**. **(F, G)** Higher LPP amplitudes were still associated with shorter RTs for each ambiguity level when we sorted trials according to RT for each individual face (each identity and each morph level). **(F)** ERP. **(G)** Mean LPP amplitude. **(H-L)** The LPP amplitude was still associated with behavioral response when using all trials (without excluding any trials with RT<1100ms). Higher LPP amplitudes were associated with shorter RTs regardless of the same stimulus, but similar LPP amplitudes were found for similar RTs regardless of different ambiguity levels. **(H)** Mean LPP amplitude.

The LPP was sorted into 4 groups according to RT for each subject. **(I, K)** ERP. **(J, L)** Mean LPP amplitude. Conventions as **Fig. 3**.

Supplementary Fig. 3. Single-trial analysis revealed a trial-by-trial coupling between the LPP and RT. **(A-D)** Two single-subject examples showing trial-by-trial relationship between the LPP amplitude and RT. **(A, C)** Single-trial ERP sorted by RT. Color coding shows the ERP amplitude. Trials are aligned at stimulus onset (Time=0; solid black line), and are sorted by RT (dashed black line). The gray line denotes stimulus offset and the dashed black rectangle denotes the LPP interval. The lower plot shows the mean ERP averaged across all trials. Higher LPP amplitudes resulted in shorter RTs. **(B, D)** Correlation between the LPP amplitude and log-transformed RT (in ms; relative to stimulus offset). Strong negative correlations ($r=-0.26$, $P=1.6\times10^{-4}$ for **(B)** and $r=-0.17$, $P=0.026$ for **(D)**) were observed for both example subjects. Each dot represents a single trial and the gray line represents the best linear fit. **(E)** Summary of single-trial correlation coefficients. Each bar represents a single subject, sorted by the correlation coefficient (from the most negative to the most positive). Significant correlations are marked in purple. Although the distribution indicates some individual differences, most subjects demonstrated a negative correlation between the LPP amplitude and RT (χ^2-test: $P=0.008$). **(F-H)** Summary of single-trial correlation coefficients when using raw RT. Most subjects still demonstrated a negative correlation between the LPP amplitude and RT (one-sample t-test against 0 on Fisher's Z-transformed *r*, all $Ps<0.05$). **(F)** Excluding trials with RT<1100ms. **(G)** All trials. **(H)** Correlation with LPP peak amplitude.

Supplementary Fig. 4. Context modulation of the LPP, separately for the condition without breaks between blocks **(A-F)** and the condition with breaks between blocks **(G-L)**. **(A, G)** Accuracy of judging unambiguous faces did not differ between blocks. **(B, H)** RT of unambiguous faces did not differ between blocks. **(C, I)** RT differed significantly as a function of ambiguity levels in the second block. **(D, J)** The LPP was not only modulated by ambiguity levels, but also by the context of ambiguous stimuli. **(E, K)** Mean LPP amplitude for unambiguous faces across blocks. LPP amplitudes were averaged across the entire interval (shaded area). **(F, L)** Mean LPP amplitude for faces with different levels of ambiguity in the second block. Error bars denote one SEM across subjects. Asterisk indicates a significant difference between conditions. +: $P<0.1$, *: $P<0.05$, **: $P<0.01$, and ***: $P<0.001$. n.s.: not significant.

Supplementary Table 1. Brain areas modulated by fear or ambiguity levels. All values are P<0.001 uncorrected. Asterisk indicates P<0.05 FWE after small volume correction.

	Brain Region	Z-score	Peak Coordinate MNI (X Y Z)	Volume (voxel)
Increasing fear	R Motor Cortex (R Precentral Gyrus)	5.96	42 −15 60	1174
	L Fusiform Face Area	4.52	−24 −48 −24	174
	L Supramarginal Gyrus	4.36	−63 −54 21	49
	R Paracentral Lobule	4.31	12 −21 51	47
Decreasing fear	L Anterior Insula	4.60	−45 0 15	221
	L Motor Cortex (L Precentral Gyrus)	5.76	−45 −21 48	1215
	L Amygdala	3.22*	−21 −6 −15	6
	R Fusiform Face Area	5.28	36 −54 −30	378
	L Occipital Cortex/BA 18	4.38	−6 −105 9	136
	L Ventral ACC	4.29	−6 3 51	64
Increasing ambiguity	L Dorsal mPFC/ACC	4.83	−6 15 54	465
	R Dorsal mPFC	3.62	57 18 39	74
	L Inferior Frontal Gyrus/ Anterior Insula	4.64	−30 27 0	215
	R Inferior Frontal Gyrus/ Anterior Insula	4.60	36 21 −6	235
	Caudate Body	4.33	12 9 15	68
Decreasing ambiguity	R Amygdala	3.17*	30 0 −21	17
	L Ventral ACC	4.91	−6 39 −9	454
	L PCC	4.29	−12 −45 42	263
	L Dorsolateral Prefrontal Cortex	4.87	−30 21 45	305
	R Postcentral Gyrus	5.87	57 −24 18	779

	L Inferior Parietal Lobule	5.02	−54 −48 48	819
	Superior Temporal Gyrus	4.50 5.50	−63 −18 −30 −66 −30 3	157 144

Supplementary Results

Logistic mixed-model regression for behavioral judgments

We used a logistic mixed model to fit behavioral judgments for all subjects with subject group and ambiguity level as the fixed effects and each subject as the random effect. We found a significant main effect of ambiguity levels (two-way mixed-measure ANOVA; $F(1,194)=6.26$, $P=0.013$, $\eta_p^2=0.21$), with a higher percentage of fearful judgments for high (53.8±1.54%; mean ±SD) and intermediate (53.7±1.00%) ambiguity but not unambiguity (50.3±0.75%), suggesting a bias towards fearful judgments for ambiguous faces. No significant difference was found between three subject groups ($F(1,194)=0.14$, $P=0.71$, $\eta_p^2=0.01$), nor a significant interaction ($F(1,194)=0.72$, $P=0.398$, $\eta_p^2=0.03$).

Confidence judgment and reaction time (RT) reflect perceptual ambiguity

For behavioral subjects, after reporting a face as fearful or happy, we asked them to report their confidence in their decisions. Subjects reported significantly higher levels of confidence for unambiguous faces compared to ambiguous faces (one-way repeated-measure ANOVA of morph levels (**Supplementary Fig. 1A1**): $F(6,138)=42.0$, $P<0.001$, $\eta_p^2=0.65$; one-way repeated-measure ANOVA of ambiguity levels (**Supplementary Fig. 1A2**): $F(2,46)=72.6$, $P<0.001$, $\eta_p^2=0.76$). In addition to the explicit confidence ratings, the RT (relative to stimulus onset) for the fear/happy decision can be considered as an implicit measure of confidence because it showed a similar pattern as the explicit ratings (**Supplementary Fig. 1A3, A4**). RT was faster for unambiguous faces compared to ambiguous faces for all three subject groups (**Fig. 1D** for all subjects; one-way repeated-measure ANOVA of morph levels, $F(6,390)=27.6$, $P<0.001$, $\eta_p^2=0.30$; **Supplementary Fig. 1A-C** for each subject group; behavioral: $F(6,138)=19.0$, $P<0.001$, $\eta_p^2=0.45$; EEG: $F(6,132)=5.57$, $P<0.001$, $\eta_p^2=0.20$; fMRI: $F(6,108)=6.27$, $P<0.001$, $\eta_p^2=0.26$). When grouping all trials into three levels of ambiguity, subjects showed shortest RT for unambiguous faces but longest RT for high ambiguity faces (**Fig. 1E** for all subjects; unambiguous: 1.42±0.16s (mean±SD), intermediate: 1.46±0.17s, and high: 1.51±0.20s; one-way repeated-measure ANOVA of ambiguity levels, $F(2,130)=47.1$, $P<0.001$, $\eta_p^2=0.42$), showing a systematic relationship between RT and ambiguity: the more ambiguous, the longer the RT. This was also the case for each individual subject group (**Supplementary Fig. 1A-C**; one-way repeated-measure ANOVA of ambiguity levels; behavioral: $F(2,46)=31.5$, $P<0.001$, $\eta_p^2=0.58$; EEG: $F(2,44)=12.5$, $P<0.001$, $\eta_p^2=0.36$; fMRI: $F(2,36)=10.5$, $P<0.001$, $\eta_p^2=0.37$). Note that the behavioral patterns of all three subject groups were comparable, although EEG and fMRI subjects did not provide confidence ratings. Furthermore, behavioral subjects judged facial emotions faster (**Supplementary Fig. 1A5**; one-way repeated-measure ANOVA of confidence levels, $F(2,46)=39.2$, $P<0.001$, $\eta_p^2=0.63$) and reported confidence faster (**Supplementary Fig. 1A6**; $F(2,46)=16.0$, $P<0.001$, $\eta_p^2=0.41$) when they subsequently indicated higher confidence. Moreover, their RT to report confidence also varied marginally significantly as a function of morph levels (**Supplementary Fig. 1A7**; $F(6,138)=2.17$, $P=0.094$, $\eta_p^2=0.09$), and varied significantly as a function of ambiguity levels (**Supplementary Fig. 1A8**; $F(2,46)=3.71$, $P=0.046$, $\eta_p^2=0.14$). Together, both explicit confidence ratings and implicit confidence measures by RT demonstrated a systematic relationship with perceptual ambiguity: the more ambiguous, the lower the confidence and the longer time it took to process the ambiguous information.

The peak amplitude but not latency of the LPP encodes ambiguity

We have shown that the mean LPP amplitude encodes ambiguity. Here, we further confirmed our results using the peak amplitude of the LPP. The peak amplitude was the most positive voltage of the entire LPP interval (400 to 700ms after stimulus onset). Similarly, the peak amplitude of the LPP varied systematically as a function of ambiguity (one-way repeated-measure ANOVA of ambiguity levels, $F(2,44)=11.1$, $P=1.25\times10^{-4}$, $\eta_p^2=0.34$), and post-hoc t-tests revealed significant differences between unambiguity (8.20±2.26μV, mean±SD) and intermediate ambiguity (7.05±1.99μv; paired two-tailed t-test, $t(22)=2.90$, $P=0.008$, $d=0.61$), and marginally significant difference between intermediate and high ambiguity (6.48±1.80μv; $t(22)=1.95$, $P=0.064$, $d=0.42$). However, the latency corresponding to the peak amplitude was similar across ambiguity levels ($F(2,44)=0.86$, $P=0.43$, $\eta_p^2=0.038$; unambiguous: 540±65ms (mean±SD), intermediate:

524±65ms, high: 521±63ms). Together with our results of the mean LPP amplitude, we show that the LPP encodes ambiguity with amplitude but not latency.

Time-frequency analysis reveals oscillations in the delta frequency band that parametrically encode ambiguity levels

We have shown that the LPP is involved in ambiguity processing, but that analysis pooled over all frequency bands. Might ambiguity processing be reflected in brain oscillations at particular frequencies? We next conducted a time-frequency analysis. Brain oscillations were strongly enhanced between 100 and 300ms in the theta (4-7Hz) power band, between 400 and 700ms in the delta (1-4Hz) power band, but reduced between 400 to 700ms in the alpha (8-12Hz) and beta (13-25Hz) bands, at the electrode Pz (**Fig. 2F**). Notably, event-related spectral perturbation (ERSP) was significantly higher for unambiguous stimuli compared to more ambiguous stimuli in the delta frequency band (**Fig. 2G-H**; one-way repeated-measure ANOVA of ambiguity levels: $F(2,44)=13.5$, $P=2.71\times10^{-5}$, $\eta_p^2=0.38$; unambiguous vs. intermediate: $t(22)=3.07$, $P=0.0056$, $d=0.65$; intermediate vs. high: $t(22)=2.42$, $P=0.024$, $d=0.51$). However, ERSP was remarkably similar for all three ambiguity conditions in the theta, alpha and beta frequency bands (one-way repeated-measure ANOVA, all $F(2,44)<1$ and all $Ps>0.05$). Similar ERSP results in the delta frequency band were derived in the speeded version of the task as well ($F(2,44)=6.16$, $P=0.0044$, $\eta_p^2=0.29$).

Control analysis for the LPP and RT

We found very similar results when we sorted trials according to RT for each individual face (each identity and each morph level) (**Supplementary Fig. 2F, G**; main effect of ambiguity level: $F(2,44)=17.4$, $P=2.77\times10^{-6}$, $\eta_p^2=0.44$; main effect of RT group: $F(1,44)=8.45$, $P=0.002$, $\eta_p^2=0.28$; interaction: $F(2,44)=1.54$, $P=0.22$, $\eta_p^2=0.07$), arguing against any effect of the specific stimulus on the LPP. Similar results were found when using the peak amplitude of the LPP (main effect of ambiguity level: $F(2,44)=12.6$, $P=4.63\times10^{-5}$, $\eta_p^2=0.37$; main effect of RT group: $F(1,44)=12.1$, $P=0.002$, $\eta_p^2=0.35$; interaction: $F(2,44)=0.56$, $P=0.58$, $\eta_p^2=0.025$).

Single-trial analysis reveals a trial-by-trial coupling between the LPP and RT

To directly capture the trial-by-trial variation in the LPP and RT, we conducted a single-trial analysis. Two example subjects showed that greater LPP amplitudes were associated with faster RTs (**Supplementary Fig. 3A-D**). To formally quantify such coupling, we correlated the mean LPP amplitude of the entire LPP interval (400 to 700ms after stimulus onset) with log-transformed RT for each subject (Pearson correlation; **Supplementary Fig. 3B**: $r=-0.26$, $P=1.6\times10^{-4}$; **Supplementary Fig. 3D**: $r=-0.17$, $P=0.026$). We found that significantly more subjects (19/23; 83%) exhibited a negative correlation between the LPP and RT (**Supplementary Fig. 3E**; χ^2-test, $P=0.008$), among which 4 subjects exhibited a negative correlation that was significant even at the level of that individual subject. Correlation coefficients (r) were significantly biased towards negative (one-sample t-test against 0 on Fisher's Z-transformed r, $t(22)=-4.54$, $P=1.62\times10^{-4}$, $d=-0.97$; Wilcoxon signed rank test on r (test for 0 median), $P=3.7\times10^{-4}$). Interestingly, individual differences in such correlations could not be explained by age (Pearson correlation with Z-transformed r, $r=-0.17$, $P=0.44$), mean LPP amplitude ($r=0.18$, $P=0.42$), nor EJI ($r=0.10$, $P=0.65$), but there was a curious correlation with the mean RT of a subject ($r=-0.42$, $P=0.046$), indicating that the LPP was more strongly associated with decisions in slower subjects.

We used log-transformed RT in this analysis, but we also found similar results using raw RT, either excluding trials with RT<1100ms (**Supplementary Fig. 3F**; $t(22)=-2.90$, $P=0.008$, $d=-0.62$), or including all trials (**Supplementary Fig. 3G**; $t(22)=-2.91$, $P=0.008$, $d=-0.62$). Notably, we found similar negative correlations with raw RT using peak amplitude of the LPP (**Supplementary Fig. 3H**; $t(22)=-3.33$, $P=0.003$, $d=-0.71$).

Together, with the highest resolution of single-trial analysis, our results further showed that greater LPP amplitudes were associated with faster behavioral judgments under ambiguity.

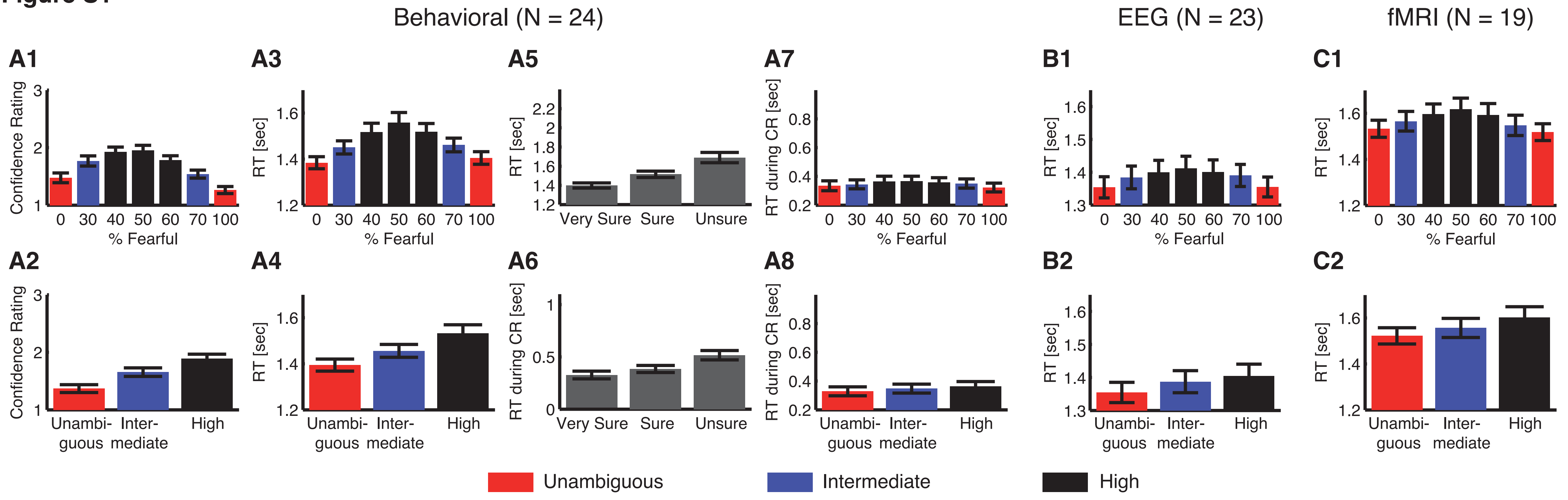
The LPP encodes domain-general ambiguity about choices

We have shown that the LPP encodes decisions about ambiguous faces. But does the LPP only encode ambiguity about faces, or even only along the fear-happy dimension? To answer this question and establish the generalizability of our findings, we conducted another two control experiments, in which 11 healthy subjects judged facial emotions along a different emotion dimension, the anger-disgust dimension (emotion task), and in which they judged animal categories of cat-dog morphs (animal task), using the stimuli from Control Experiment 2 (**Fig. 1B**). Indeed, in both control tasks, we not only found a clear LPP component in the interval of 400 to 700ms after stimulus onset at the electrode Pz as we had found in the fear-happy morph task, but importantly, the LPP also differentiated the ambiguity levels, once again with the most unambiguous stimuli showing the most positive potential. This observation was confirmed by the mean LPP amplitude (emotion task: unambiguous: 6.03±3.70μV (mean±SD), intermediate: 4.09±2.87μV, high: 3.30±1.94μV; one-way repeated-measure ANOVA of ambiguity levels, $F(2,20)=10.3$, $P=0.001$, $\eta_p^2=0.51$; animal task: unambiguous: 4.84±2.89μV, intermediate: 4.22±2.59μV, high: 2.85±2.40μV; $F(2,20)=12.8$, $P=2.62\times10^{-4}$, $\eta_p^2=0.56$). Together, our results suggest that the LPP encodes ambiguity in general, and is not limited to processing facial emotions or even faces.

Control analysis for context modulation of the LPP

Our results were further confirmed using the peak amplitude of the LPP at Pz (one-way repeated-measure ANOVA of ambiguity levels in the second block: $F(2,62)=11.3$, $P=6.81\times10^{-5}$, $\eta_p^2=0.266$; one-way repeated-measure ANOVA of block for unambiguous faces: $F(2,62)=2.90$, $P=0.060$, $\eta_p^2=0.086$). No significant difference was found on the peak latency of the LPP (all $Ps>0.05$). In addition, similar results were also observed when using the average of 6 electrodes surrounding Pz (CP3, CPz, CP4, P3, Pz and P4) for both mean (one-way repeated-measure ANOVA of ambiguity levels in the second block: $F(2,62)=25.5$, $P=8.18\times10^{-9}$, $\eta_p^2=0.45$; one-way repeated-measure ANOVA of block for unambiguous faces: $F(2,62)=6.25$, $P=0.0030$, $\eta_p^2=0.17$) and peak amplitude of the LPP (one-way repeated-measure ANOVA of ambiguity levels in the second block: $F(2,62)=12.9$, $P=2.03\times10^{-5}$, $\eta_p^2=0.29$; one-way repeated-measure ANOVA of block for unambiguous faces: $F(2,62)=3.00$, $P=0.057$, $\eta_p^2=0.088$). Again, no significant difference was found on the peak latency of the LPP (all $Ps>0.05$).

Figure S1



Behavioral (N = 24)
EEG (N = 23)
fMRI (N = 19)
A1
Confidence Rating
% Fearful
A2
Confidence Rating
Unambi- guous
Inter- mediate
High
A3
RT [sec]
% Fearful
A4
RT [sec]
A5
RT [sec]
Very Sure
Sure
Unsure
A6
RT during CR [sec]
A7
RT during CR [sec]
% Fearful
A8
RT during CR [sec]
B1
RT [sec]
% Fearful
B2
RT [sec]
C1
RT [sec]
% Fearful
C2
RT [sec]
Unambiguous
Intermediate
High

Figure S2

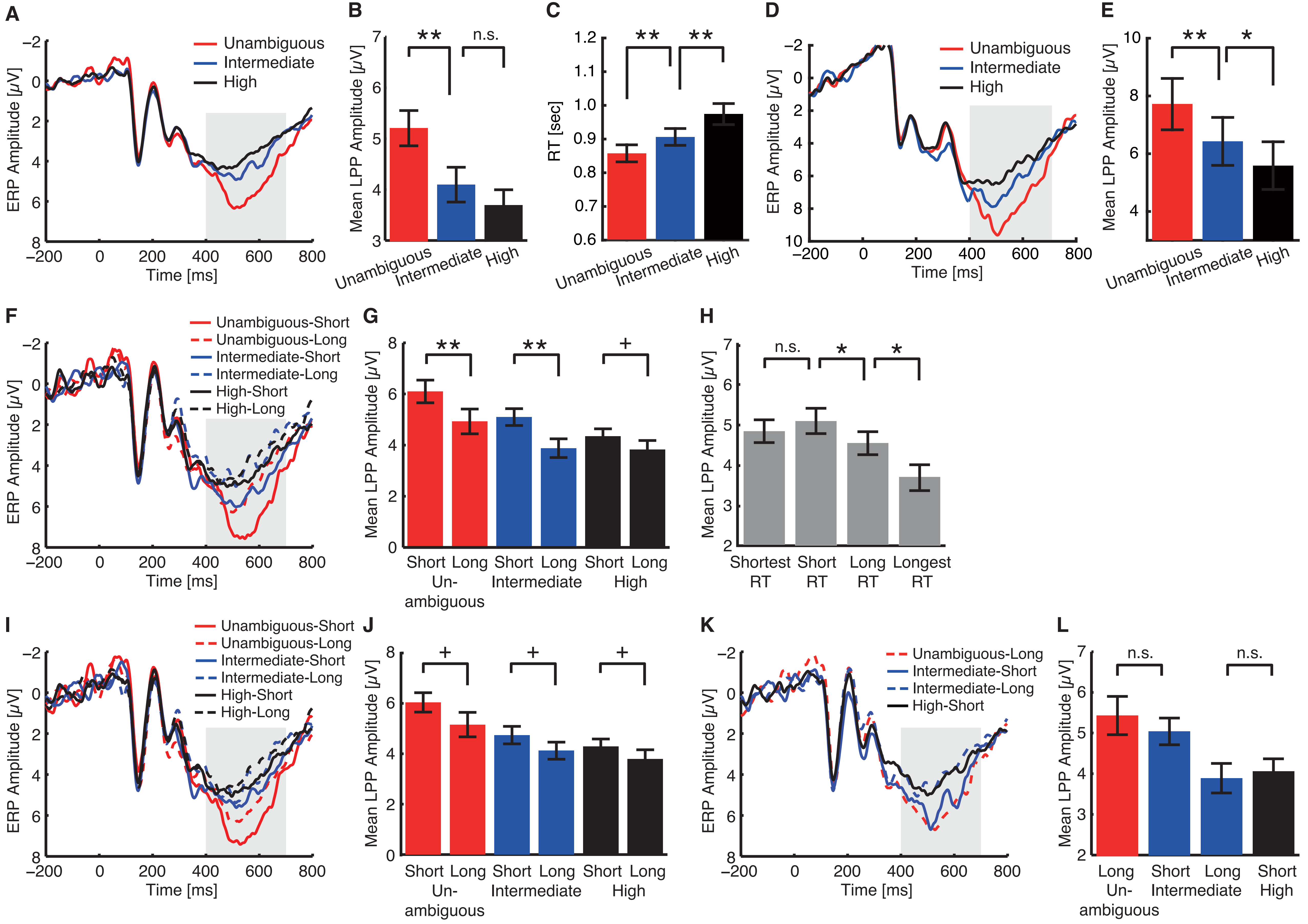

A
Unambiguous
Intermediate
High
ERP Amplitude [μV]
Time [ms]
B
Mean LPP Amplitude [μV]
**
n.s.
Unambiguous
Intermediate
High
C
RT [sec]
**
**
D
E
*
F
Unambiguous-Short
Unambiguous-Long
Intermediate-Short
Intermediate-Long
High-Short
High-Long
G
+
Short
Long
Un-ambiguous
Intermediate
High
H
Shortest RT
Short RT
Long RT
Longest RT
I
J
K
L

Figure S3

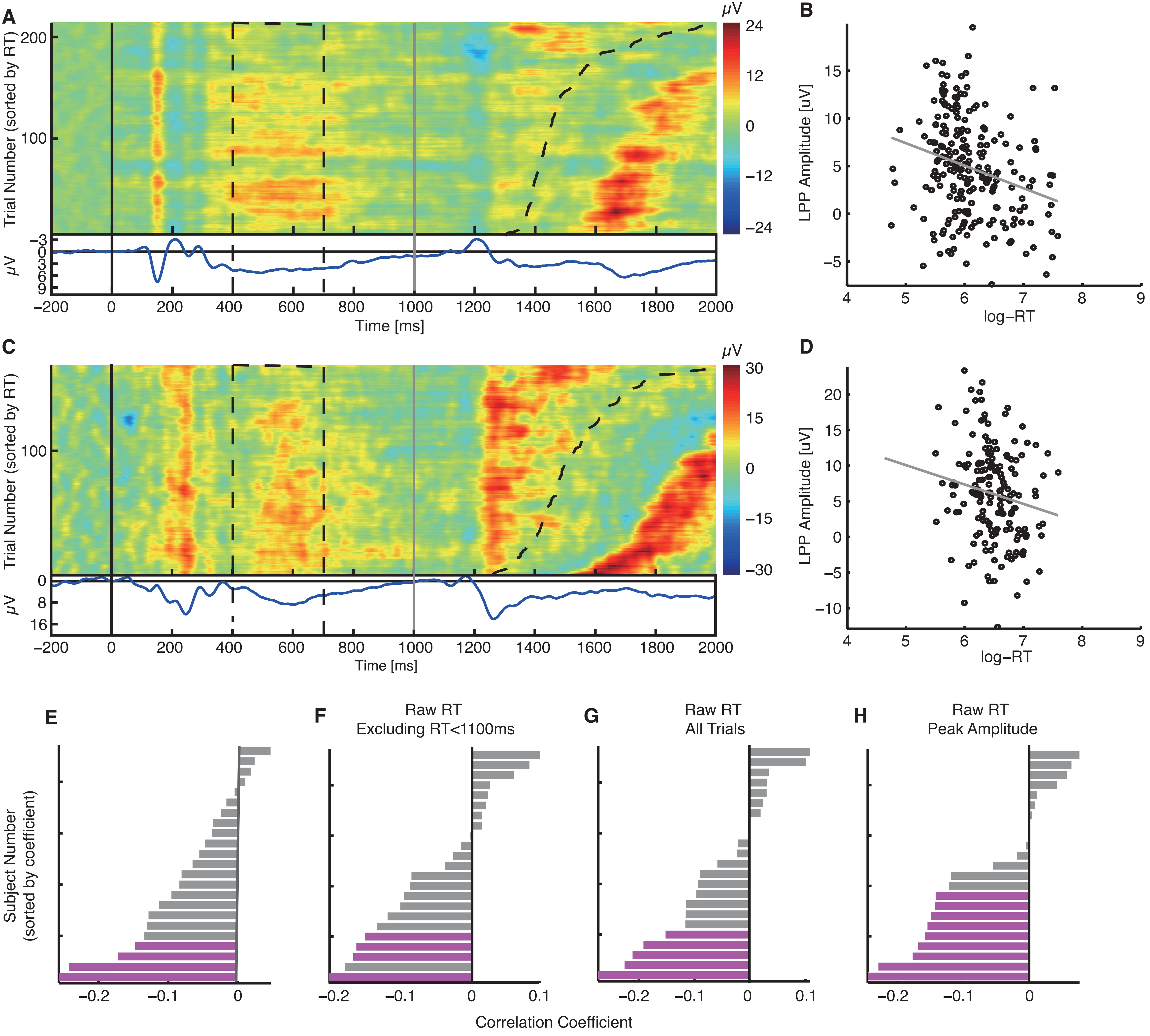
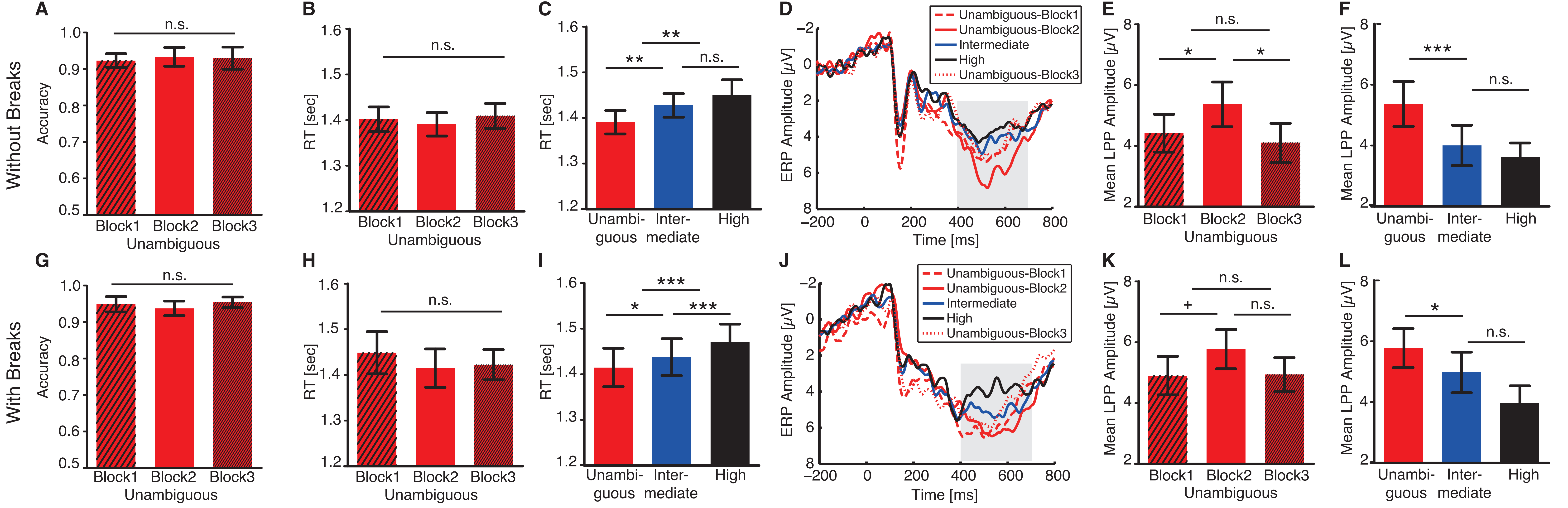

A
B
C
D
E
F
G
H
Trial Number (sorted by RT)
μV
Time [ms]
LPP Amplitude [uV]
log−RT
Raw RT
Excluding RT<1100ms
Raw RT
All Trials
Raw RT
Peak Amplitude
Subject Number
(sorted by coefficient)
Correlation Coefficient

Figure S4



A
B
C
D
E
F
G
H
I
J
K
L
Without Breaks
With Breaks
Accuracy
RT [sec]
ERP Amplitude [μV]
Mean LPP Amplitude [μV]
n.s.
**
***
*
+
Block1
Block2
Block3
Unambiguous
Unambi-guous
Inter-mediate
High
Time [ms]
Unambiguous-Block1
Unambiguous-Block2
Intermediate
High
Unambiguous-Block3